KINECT Face Recognition Using Occluded Area Localization Method

Fatema Tuz Zohra(✉) and Marina Gavrilova

Department of Computer Science, Faculty of Science, University of Calgary,
2500 University Drive N.W., Calgary, AB T2N 1N4, Canada
fatematuz.zohra@ucalgary.ca, marina@cpsc.ucalgary.ca

Abstract. Automated face recognition is commonly used for security reinforcement and identity verification purposes. While significant advancement has been made in this domain, modern surveillance techniques are still dependent on variations in pose, orientation of the facial images, difference in the illumination, occlusion, etc. Therefore, face recognition or identification in uncontrolled situations has become an important research topic. In this paper, we propose a new face recognition technique that takes into account partial occlusion, while still accurately identifying the user. The occluded facial areas are detected from the Kinect depth images by extracting features using Uniform Local Binary Pattern (LBP). For localizing occluded regions from the Kinect depth images, a threshold based approach is used to identify the areas close to the camera. The recognition system will discard the occluded regions of the facial images and match only the non-occluded facial part with the gallery of images to find the best possible match. The performance of the recognition system has been evaluated on EUROKOM Kinect face database containing different types of occluded and non-occluded faces with neutral expressions. Experimental results show that the proposed method improves the recognition rate by 4.8% and 5.7% for occlusion by hand and occlusion by paper, respectively.

Keywords: Kinect sensor · Depth images · Binary classifier · KNN classifier · Occlusion detection · Face recognition using depth information

1 Introduction

Automatic face recognition has established itself as a key research area in computer vision and pattern recognition over the past few decades. Face biometric based identification is one of the most popular and highly accepted biometric traits due to its non-invasive nature of the acquisition process. Over the years, face biometric has been extensively investigated to improve the recognition performance. However, despite many years of research, face biometric recognition is still an active domain due to the challenges in identifying the faces in unconstrained environments. These include different illumination, expression and pose variance, deformation due to aging, and different types of occlusions [4,11,15,19].

M.L. Gavrilova et al. (Eds.): Trans. on Comput. Sci. XXX, LNCS 10560, pp. 12–28, 2017.
https://doi.org/10.1007/978-3-662-56006-8_2

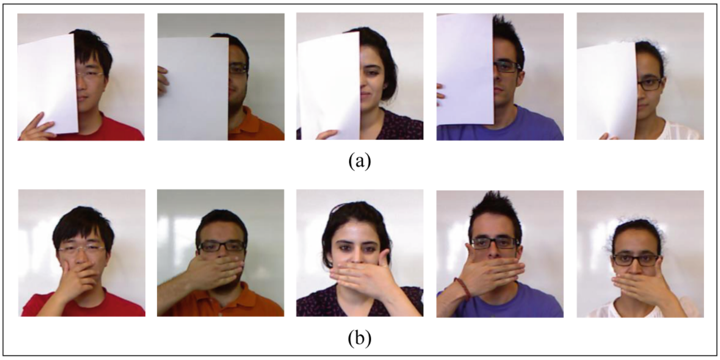


Fig. 1. Example of occluded RGB facial images from EURECOM Kinect Face dataset [20]. (a) occlusion by paper, and (b) occlusion by hand. (Color figure online)

Handling occlusions during face automation of security systems is a challenging task. Due to occlusions, different facial parts are not visible and also inaccessible during face recognition, which make it difficult to effectively recognize faces in the presence of occlusion. Sometimes it is not possible to extract distinctive facial information from the face in the presence of occlusion, which may result in false identification of the person. It should be treated carefully as it can undermine the correct performance of a surveillance system. Some recent research shows that, face recognition systems that use 3D or 2.5D (depth) information have more efficient recognition than the 2D based face recognition systems [2,3,7,13,20,23]. This is due to the fact that illumination invariant 3D information, such as depth information and a shape of a face can be incorporated with the 2D information to improve the accuracy of the recognition. However, these recognition methods also generate inferior results for facial expression and occlusion variations. Therefore, it is necessary to build a system that can effectively recognize faces in the presence of occlusion. To address this challenge, we propose a new face recognition system that will take into account the occluded area of the faces while identifying the user.

Recently, RGB-D cameras such as Kinect sensor have received a vast amount of attention from diverse research communities [12,14,16,26,27] as it is a low cost device which can effectively extract the depth mapping from the object in front of the camera. Kinect sensor can capture 2D and 3D data simultaneously with a promising acquisition time. Therefore, researchers investigated the Kinect Face datasets [20] for the purpose of face recognition, face salient points localization, occlusion detection, 3D modeling of the faces and emotion recognition. There is very limited research that concentrates on face recognition in the presence of occlusion using the depth information. Therefore, in this paper, we propose a novel face recognition system that will consider the occluded area localized from Kinect depth images while identifying the user from the gallery facial images. If the probe image contains occluded area, then the face recognition system will match only the non-occluded facial parts with the gallery images to find the best possible match. This will reduce the number of misclassification, and as a result, will improve the recognition performance of the biometric system. We extract the local features from the facial images using Local Binary Pattern (LBP) analysis and feed those features to the k-nearest neighbor (KNN) classifier to identify the occluded faces. For detecting and localizing the occluded facial areas we used the depth information provided by Kinect RGB-D camera. For the evaluation of the proposed method, we consider EURECOM Kinect Face dataset [20]. This journal paper is an extended version of the research presented in Cyberworlds 2016 [34]. Figure 1 shows an example of occluded RGB facial images from the EURECOM Kinect Face dataset.

The rest of the paper is structured as follows. Section 2 presents a comprehensive study of the previous research works on Kinect database and face occlusion detection approaches based on depth information. In Sect. 3 detailed description of the proposed method for face recognition using occluded area localized from Kinect depth images is presented. The experimental setup and results are shown in Sect. 4. Finally, Sect. 5 concludes the discussion with some future directions.

2 Literature Review

Occlusion detection has been extensively studied in the literature. Many approaches have been proposed to detect 2D faces under occlusions. Some of the very recent approaches [10,18,19,33] can accurately detect the occlusion in 2D facial images. An effective method for occlusion detection based on the structured sparse representation is used to estimate the error introduced by occlusion in [18]. The authors used a morphological graph model and localized similarity measure to detect the error. The proposed method can handle high level of occlusion. In [10], the authors proposed an approach to detect half occluded faces in an unconstrained crowd scene. It used a machine learning approach and trained the classifier with left and right half occluded faces. The authors proposed a free rectangular feature to modify the Viola-Jones algorithm for detecting the occluded faces. They also applied skin color models to improve the correctness of the system. The proposed method is evaluated on FDDB dataset. The first occlusion detection method based on multi-task convolution neural network (CNN) was proposed in [33]. The authors pre-trained the CNN model with non-occluded facial images of different orientations. After that, they fine-tuned all the convolution filter using the deep FO datasets and determined the occluded facial parts. The proposed method shows accuracy of 100% and 97.24% while evaluating on the AR face dataset and FO dataset respectively.

Recently, researchers have focused on face recognition based on 2.5D and 3D facial images [2,3,7,13,20,23]. However, there is very limited research on occlusion detection and localization based on depth images. In [5], the authors proposed an automated approach that can detect, normalize and recognize faces in the presence of extraneous objects from depth images. They built a non-occluded face model using eigenfaces approach and compared every part of the face with this model. Any part that did not match with the model was identified as the occluded facial part. They tested their proposed approach on UND database processed with an artificial occlusion generator. Their method could detect 83.8% of the occluded faces. The main deficiency of this approach is that it used artificially generated occlusions which may not reflect the real environment. The authors of [1] proposed an occlusion detection method for unconstrained three-dimensional (3D) facial identification. They used a nose-based registration of the facial image to align the input face with a generic face model. The occluded areas are detected by computing the absolute difference between the input face and the generic face model. For evaluation, they considered Bosphorus 3D face database resulting in an accuracy of 94.23%. For the nose-based registration to work, the nose area should be visible. Otherwise, the proposed method will fail. In [28], the authors proposed a similar approach as [5] for occlusion detection based on 3D face restoration. They used Gappy Principal Component Analysis (GPCA) for the restoration process. Two subsequent steps- initial 3D occlusion detection and refined 3D occlusion calculation - are used to determine the location of the occlusion. The proposed method generates an average recognition accuracy of 93.03% on UMB-DB database. In the recent paper on occlusion detection and localization [9], the authors proposed a threshold and block based approach using depth images. The depth images are processed to detect the outward sections of the face using Energy Range Face Images to identify the occluded region. The accuracy of the threshold based technique is 91.79% and the accuracy of the block based approach is 99.71% on Bosphorus database. In this method, facial image with two outward regions is identified as the occluded facial image, where one of the outward regions is nose and another one is the occluded area. However, this method will not work for the input images where the nose area is occluded. In [8], the authors proposed an occlusion detection approach that projects the facial parts into two local subspaces: PCA and two-dimensional LDA. The authors of [8] used Viola Jones algorithm to separate every facial parts and in the classification stage individually recognized each part. They used a threshold value to classify occluded and non-occluded facial parts. The proposed method was evaluated on three different databases. The accuracy of occlusion detection is 62.97% (for 2DPCA). This is the only work that considered the Kinect face dataset [20] for occlusion detection. However, they used the 2D facial images captured using Kinect sensor.

All these papers focus on occlusion detection based on 2.5D and 3D face information acquired using 3D scanners, which requires relatively high data acquisition time. However, for an effective face automation and security surveillance the data acquisition time should be lower. Therefore, in this paper, we focus on facial data captured using Kinect sensor due to its low data acquisition time. We use an occlusion detection method based on Kinect depth information. The occluded regions are located in the facial images based on the depth information. The proposed face recognition technique will utilize this depth information while identifying faces under partial occlusions. The occluded facial parts will be ignored in the recognition technique for improving the performance.

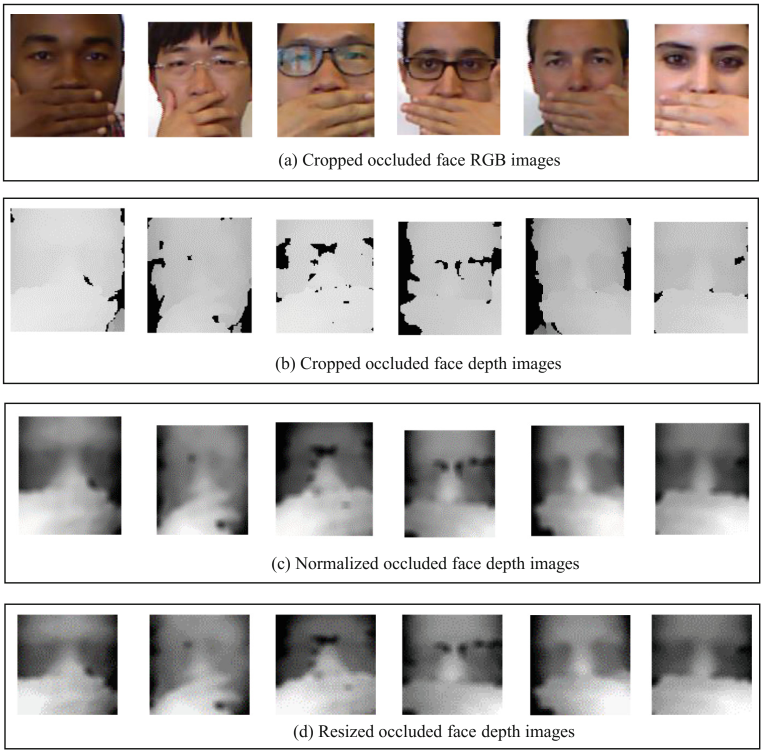
3 Methodology

Human visual system has the perception of depth. But depth information is absent in the 2D imaging system. Existing research that utilized 2.5D or 3D images for face recognition or occlusion detection uses relatively slow 3D scanners [17,24]. Kinect sensor has the RGB camera to capture RGB images, as well as an infrared (IR) laser emitter and an IR camera to act as a depth sensor [20]. Thus, Kinect sensor can capture the 2D and 3D data simultaneously with a promising acquisition time. It can build an effective depth map with varying intensity values based on the distance from the camera. The high intensity value means the object is closer to the camera and low intensity depicts the object is further from the camera. In this paper, we proposed a face recognition method that will consider the depth information acquired using the Kinect RGB-D camera for identifying faces in the presence of occlusion. The proposed method has four steps, namely: (i) preprocessing, (ii) occlusion detection, (iii) occlusion localization and (iv) face recognition under occlusion. We use LBP operators to analyze the local patterns and SVM as a binary classifier to determine whether it is a front face or an occluded face. After that, occlusions are localized in those detected occluded facial images. The face recognition system considers only the non-occluded area while identifying the user from the occluded facial images.

3.1 Preprocessing

The Kinect face database contains manually annotated facial salient points, such as left eye center, right eye center, nose-tip, left mouth corner, right mouth corner and chin. The first step in preprocessing is to crop the facial images using those six anchor points on a face. These facial points can also be extracted using Viola-Jones algorithm [32]. Figure 2(a) and (b) show example of cropped occluded face RGB and depth images, respectively.

The next step is to normalize the depth values in the cropped facial image. For that, we consider only the average ± maximum gray values (consider other values as zero), and scale it into the range 0 to 1. Figure 2(c) shows an example of normalized depth images. In a depth image, intensity values indicate estimated distances from the camera [25]. The facial images are normalized to emphasize the difference in depth for different facial parts. Figure 3(a) shows an example of the front face image before normalization and Fig. 3(b) depicts an example of front face image after normalization. Here, x and y axis represent the size of the image, and z axis represents the depth values. From the figure we can see that the differences in depth for different facial parts is more visible in the normalized version of facial images. Also, we can see that intensity values indicate estimated distances from the camera. In Fig. 3(b) red pixels represent the object closer to the camera. If there is any occlusion presents in the facial image, then it will have lower distance from the camera which can be estimated from the pixel values. Figure 4(a) and (b) show examples of normalized depth images of a face occluded by hand and paper respectively. Here, x and y axis represent the size of the image, and z axis represents the depth values. The marked area



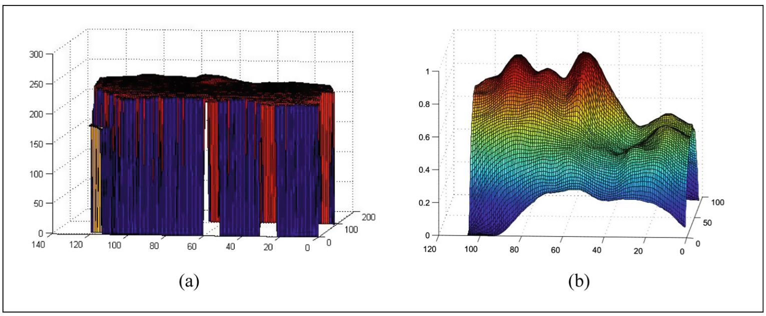
(a) Cropped occluded face RGB images

(b) Cropped occluded face depth images

(c) Normalized occluded face depth images

(d) Resized occluded face depth images

Fig. 2. Example of occluded (occlusion by hand) facial images from EURECOM Kinect Face dataset [20]. (a) cropped RGB images, (b) cropped depth images, (c) normalized depth images, and (d) resized depth images. (Color figure online)



(a) (b)

Fig. 3. Example of front face image plotted using Matlab *surface plot* operation (a) before normalization, and (b) after normalization. (Color figure online)

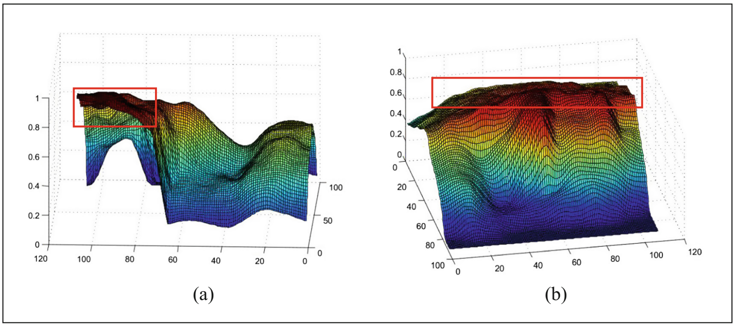


Fig. 4. Example of depth images after preprocessing. The marked area in red pixels depicts occlusion by (a) hand and (b) paper. (Color figure online)

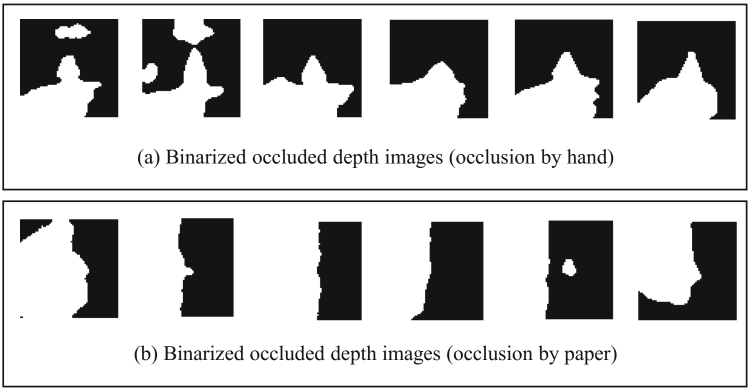


Fig. 5. Example of occluded facial images after thresholding and noise removal, (a) occluded by hand, and (b) occluded by paper.

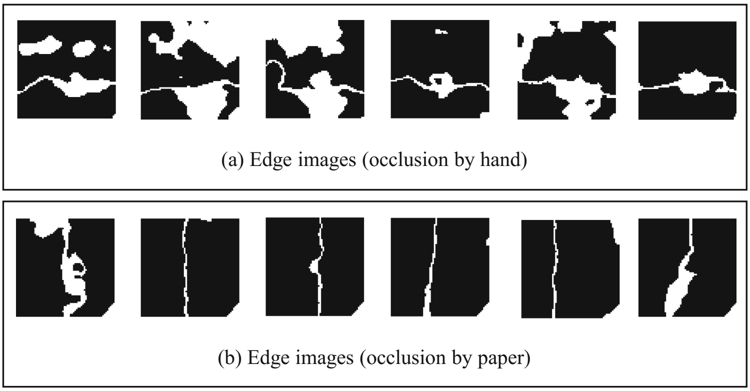


Fig. 6. Example of edge images computed from the absolute difference of reference and occluded depth images, (a) occluded by hand, and (b) occluded by paper.

in Fig. 4(a) and (b) represents the position of hand and paper in front of face. From the figure we can see that the occluded regions have different pixel values than the other facial parts and are identifiable from the facial image. The last step in preprocessing is to down-sample the facial image in a 64×64 region using bicubic interpolation. Figure 2(d) shows an example of resized depth images.

3.2 Occlusion Detection

After preprocessing of the facial images, features are extracted from the normalized and resized images using uniform local binary patterns (LBP) [21,22]. LBP is a powerful operator to analyze the local pattern. It generates a binary number by thresholding the neighbors of each pixel for labeling the pixels in the image. LBP is used extensively in local pattern analysis for its discriminating power and computational simplicity. In [20], the authors showed that for analyzing depth images, LBP works better than other feature extraction operators, such as Principle Component Analysis (PCA), Scale-invariant feature transform (SIFT) and Local Gabor binary pattern (LGBP). Thus, for our proposed method we use uniform Local Binary Patterns for feature extraction. The notation for uniform LBP is $LBP^{u2}_{(P,R)}$ where P is the number of neighbors on a circle of radius R [21,22]. The occurrences of the uniform pattern codes are extracted from the image and placed into a histogram. For our proposed method, we extract the feature histogram for different neighborhood and radius, and examine which LBP operator performs better. The investigation shows that among $LBP^{u2}_{(8,1)}$, $LBP^{u2}_{(8,2)}$ and $LBP^{u2}_{(16,2)}$ operators, $LBP^{u2}_{(8,2)}$ results in better classification accuracy for detecting front and occluded faces. Therefore, to better capture the difference between the front face and occluded face from depth images, we derive a facial representation using $LBP^{u2}_{(8,2)}$ operator which yields a feature vector of length 59.

After computing the feature histogram from the facial images using $LBP^{u2}_{(8,2)}$ operator, we use a binary classifier, support vector machine (SVM) [6] to determine whether the input image corresponds to a frontal view of the face or it contains some kind of occlusions. The classifier is first trained using a set of positive (front faces) and negative (occluded faces) samples for the classification task. In the training phase, the front face images are labeled as binary 1 and occluded facial images (occluded by hand and paper) as binary 0. The computed feature histograms along with the front face-occluded face labeling are fed into the SVM classifier to train the model.

3.3 Occlusion Localization

Investigation on depth images shows that if the face is occluded, then there must be some regions other than the nose area in the face that is closer to the camera. From Fig. 2(d), we can see that the occluded region of the face (occlusion by hand) has higher pixel values in the depth map. The pixel intensity is disproportional to the distance from the camera. Based on this hypothesis, a threshold based approach is proposed to extract high intensity values from the occluded depth images. The facial images are filtered based on empirically set threshold value, T. After that morphological opening operation is applied to the threshold image to remove all the small objects from binary images. In image processing, morphological open operation can be defined as an erosion followed by a dilation operation. It uses a structuring element for both operations. The erosion operation slides the structuring element over an image to find the local minima, and it creates the output matrix from these minimum values. If the neighborhood or structuring element has a center element then the minima is placed in that place. Similarly, in dilation operation the structuring element is rotated 180°. Then the structuring element is slid over an image to find the local maxima, and to create the output matrix from these maximum values.

Figure 5 shows an example of occluded facial images (occluded by hand and paper) after the thresholding and noise removal step (i.e. morphological opening). In the next step, a component with the maximum energy (i.e. highest pixel intensities) is selected as the potential candidate for the occluded region after connected component analysis on the binary image. Connected-component analysis is an algorithmic application of graph theory. In this analysis, subsets of connected components are uniquely labeled based on a given heuristic. The selected occluded area is then corrected using the reference front face image. In creating the reference image, we have considered the 200 front face images from the database. The absolute difference between the reference image and the occluded facial image will result in an image that has higher pixel values in the area where the difference between the reference and occluded facial image is higher. The resulting images are then binarized to get images with edges at the boundary of the occluded region. In our proposed method, we referred to these images as edge images. Figure 6 shows the example of edge images. In the next step, the selected occluded area after connected component analysis is corrected using the edge images. The edge image contains the boundary of the occluded area. Based on this boundary, the connected component is adjusted to find the accurate occluded area of the facial image. Figure 7 depicts an example of

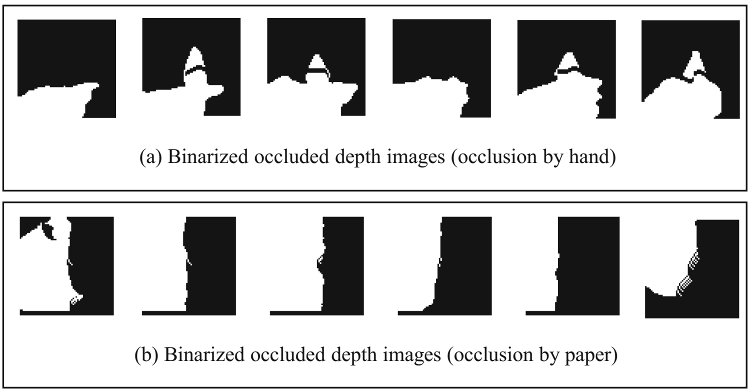


Fig. 7. Example of occluded facial images after correcting the boundary using edge images, (a) occluded by hand, and (b) occluded by paper.

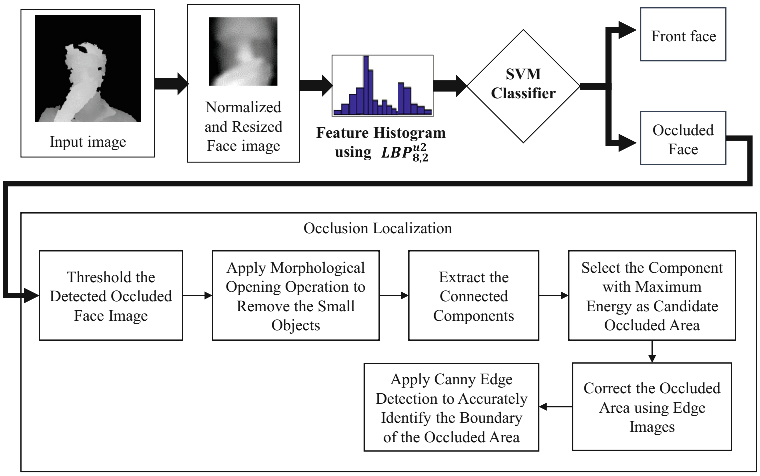


Fig. 8. Steps for occlusion detection and localization from depth images acquired using Kinect.

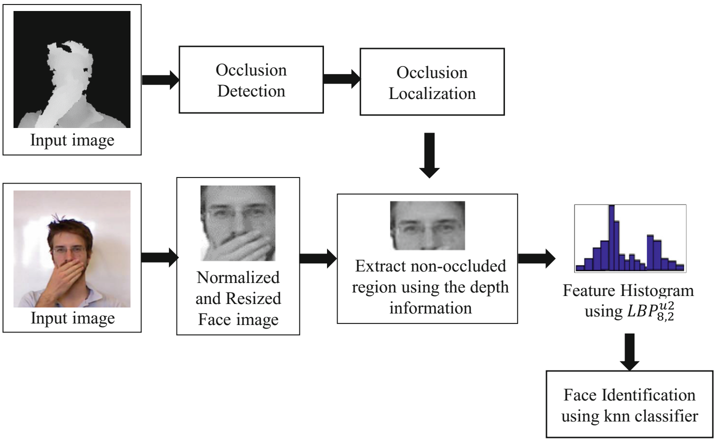


Fig. 9. Generalized block diagram for face recognition in the presence of occlusion.

binarized occluded facial images (occluded by hand and paper) after correcting the occluded region. Finally, we apply canny edge detection to fine tune the boundary of the occluded region. The canny edge detection algorithm can identify a wide range of edges in images. Figure 8 shows the steps for occlusion detection and localization from depth images.

3.4 Face Recognition Under Occlusion

For face recognition using a test image, we first detect whether the facial image is occluded or not. If there is no occlusion present in the facial image, then features are extracted from the entire image to identify the user from the gallery images. Otherwise, occlusions are localized in the facial image and non-occluded facial parts are extracted from the occluded facial image. To extract more accurate local textures of the face, the facial images are divided into 8×8 non-overlapping regions. We derive a facial representation using $LBP^{u2}_{(8,2)}$ operator from those non-overlapping regions. The feature vectors from each region are concatenated and fed to a KNN classifier for determining the face recognition performance. For simplicity of the implementation, we consider KNN for the classification task. Figure 9 shows the block diagram for face recognition in the presence of occlusion.

4 Experimental Results and Discussions

Extensive experimentation has been performed to determine and localize the occluded region in the facial images, and to determine the performance of the face recognition system using the depth information. In this section, we discuss about the experimental settings and results, and the image database used for validation of the proposed method.

4.1 Experimental Setup

To validate the proposed method, we considered the EURECOM Kinect Face Dataset [20] which is composed of 52 subjects: 14 females and 38 males. The images are captured in two sessions and there are nine types of variations in the images: neutral, smiling, open mouth, illumination variation, occlusion of half of the face by paper, occlusion of the mouth by hand, occlusion of the eyes by glasses, and left and right profile of the facial images. The database contains three different sources of information: depth bitmap image, RGB image, and 3D object files. The database also contains manual annotations of the facial salient points. From these sets of images, we considered neutral, illumination variation, occlusion of half of the face and occlusion of the mouth for experimental purpose.

For classifying the images in occluded and non-occluded classes, we used a nonlinear support vector machine (SVM) [6,31] classifier with radial basis kernel to classify front face and occluded face. SVM classifier has a simple geometric interpretation and gives a sparse solution. Moreover, the computational complexity of SVM does not depend on the dimensionality of the input space and it is less prone to over-fitting. Considering all these issues and for the simplicity of implementation, we considered SVM as the classifier for detecting occlusion. The value of sigma for SVM was set empirically. For simplicity and reproduction of the code, other existing parameters were set to their default values according to Matlab function library. For the generalization of the implemented method, 5-fold cross-validation was applied where the training fold contains images of 160 front and 160 occluded faces and the testing fold contains images of 40 front and 40 occluded faces from the database. For the recognition task, K-Nearest Neighbor (KNN) classifier is used with the distance defined as *'cityblock'*. The neutral facial images from session 1 and session 2 were used as gallery images. $LBP^{u2}_{(8,2)}$ operator is used for extracting features from 8×8 non-overlapping regions of the facial images.

4.2 Results and Discussions

For detecting occluded facial images, we consider neutral, light on, occlusion of mouth by hand, and occlusion of face by paper from the database. At first we extracted the individual feature histogram using $LBP^{u2}_{(8,1)}$, $LBP^{u2}_{(8,2)}$ and $LBP^{u2}_{(16,2)}$ operators. We investigated the accuracy of classification using individual LBP operators to find the best features for efficient classification. Figure 10 shows the plotting of average accuracy of the SVM classifier for determining front and occluded face using different LBP operators. From the graph, it is clear that $LBP^{u2}_{(8,2)}$ operator results in a higher accuracy than other LBP operators, and sigma value of 6.0 gives highest classification accuracy. Therefore, we extracted features using $LBP^{u2}_{(8,2)}$ operators. Table 1 shows the classification rate (%) using different LBP operators with different values of sigma, for identifying the front face and occluded face. In our work the occlusion of eyes by glasses was not considered. As the Kinect depth information is low quality sometimes it is difficult to extract the accurate depth information from the Kinect depth images. Thus, considering only the depth information it is quite difficult to identify the occlusion of eyes with glasses. From the detected occluded facial images, occluded region was localized using our proposed method in [34]. Figure 11 depicts an example of a localized occluded area in depth images (occluded by hand and paper) after correcting the boundary of the occluded region. After localizing the occluded region using our proposed method we visually inspected the localized area. Figure 12 shows the localized region of occlusion (occluded by hand and paper) in the RGB facial images.

Table 1. Classification rate (%) using different LBP operators.

Operator	Feature vector length	Classification rate (%)	Value of sigma
$LBP^{u2}_{(8,1)}$	59	97.25	4.5
$LBP^{u2}_{(8,2)}$	59	98.50	6.0
$LBP^{u2}_{(16,2)}$	243	96.75	9.0

After localizing the occluded area, we extracted the non-occluded region from the facial images. Features are extracted from these non-occluded regions using

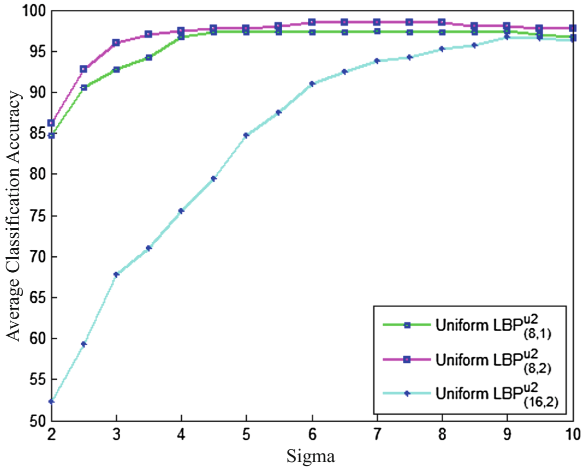


Fig. 10. Plotting of accuracy for different LBP operators.

$LBP^{u2}_{(8,2)}$ operator. The feature vectors are then fed to a KNN classifier for determining the face recognition performance. For simplicity of the implementation, we consider KNN for the recognition task. Table 2 shows the Rank-1 identification rate for the 2D face recognition in the presence of occlusion. The images contain occlusion of mouth by hand and occlusion of face by paper. From the table, we can see that the average identification rate for occlusion of mouth by hand images from the two sessions is 90.39% and the average identification rate for occlusion of face by paper is 83.66%. Table 3 shows the Rank-1 identification rate for the 2D face recognition in the presence of occlusion using our proposed method. For the images containing occlusion of mouth, the average identification rate improves to 95.19%, and for the images containing occlusion of face by paper, the average identification rate improves to 89.42%. Therefore, the proposed face recognition technique improves the recognition performance by using the depth information from the Kinect depth images. The proposed method exploits the depth information for localizing the occluded area in the facial images and discard the occluded area while matching the probe images to the gallery images.

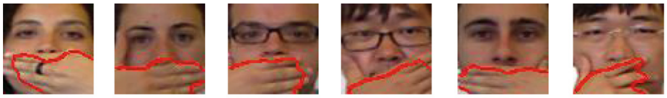
Table 2. Rank-1 identification rate for 2D face recognition under occlusion.

Session	Occlusion by hand	Occlusion by paper
Session 1	92.31%	84.62%
Session 2	88.46%	82.69%

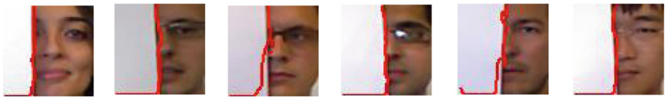
(a) Localized occluded region (occlusion by hand)

(b) Localized occluded region (occlusion by paper)

Fig. 11. Example of localized occluded area in the depth images (a) occluded by hand, and (b) occluded by paper.



(a) Localized occluded region in the RGB face images (occlusion by hand)



(b) Localized occluded region in the RGB face images (occlusion by paper)

Fig. 12. Example of localized occluded area in the RGB facial images (red marked area), (a) occluded by hand, and (b) occluded by paper (the resolution of the images are 64×64). (Color figure online)

Table 3. Rank-1 identification rate For 2D face recognition using proposed method under occlusion.

Session	Occlusion by hand	Occlusion by paper
Session 1	96.15%	90.38%
Session 2	94.23%	88.46%

5 Conclusion

Face recognition under occlusion is one of the active research topics in face automation and security surveillance. In this paper, we present a face recognition technique that will use the depth information acquired using Kinect RGB-D camera for localizing the occluded region from the facial images and use only the non-occluded regions while identifying the user from the gallery images. Local features are extracted from the depth images using LBP operators and a nonlinear SVM classifier is used to detect the front face and occluded face. For localizing occluded regions in the facial image, a threshold based approach is presented to identify the occluded area. The proposed method has been investigated on EUROKOM Kinect face database consisting of different types of occlusions and neutral facial images [20]. Experimental results show that occluded facial images can be effectively detected from the depth images, and also we can localize the occluded regions of the facial images by using the depth information. The recognition system discards the occluded regions of the facial images and match only the non-occluded facial part with the gallery images to find the best possible match. Therefore, the proposed method improves the recognition performance in the presence of occlusion in the facial images. In the future, we will consider other types of occlusions, such as occlusion induced by wearing glasses, hat or other accessories, head rotation and occlusion of different facial parts. We will also estimate the quality of the facial image under occlusion by determining the proportion of the occluded region in the facial images. Based on the ratio of the occluded area, a quality score can be assigned to the facial image. Then, the quality score can be fused with the matching score at score level fusion to determine the confidence of the facial recognition system under occlusion [29,30].

Acknowledgments. We would like to acknowledge NSERC Discovery Grant RT731064, as well as NSERC ENGAGE and URGC for partial funding of this project. Our thanks to all the members of BTLab, Department of Computer Science, University of Calgary, Calgary, AB, Canada for providing their valuable suggestions and feedback.

References

1. Alyüz, N., Gökberk, B., Spreeuwers, L., Veldhuis, R., Akarun, L.: Robust 3D face recognition in the presence of realistic occlusions. In: 5th IAPR IEEE International Conference on Biometrics (ICB), pp. 111–118, March 2012
2. Canavan, S., Liu, P., Zhang, X., Yin, L.: Landmark localization on 3D/4D range data using a shape index-based statistical shape model with global and local constraints. Comput. Vis. Image Underst. **139**, 136–148 (2015)
3. Cardia Neto, J.B., Marana, A.N.: 3DLBP and HAOG fusion for face recognition utilizing Kinect as a 3D scanner. In: Proceedings of the 30th Annual ACM Symposium on Applied Computing, pp. 66–73. ACM, April 2015
4. Chen, Y.C., Patel, V.M., Phillips, P.J., Chellappa, R.: Dictionary-based face and person recognition from unconstrained video. IEEE Access **3**, 1783–1798 (2015)
5. Colombo, A., Cusano, C., Schettini, R.: Recognizing faces in 3d images even in presence of occlusions. In: 2nd IEEE International Conference on Biometrics: Theory, Applications and Systems, BTAS, pp. 1–6, September 2008
6. Cortes, C., Vapnik, V.: Support-vector networks. Mach. Learn. **20**(3), 273–297 (1995)
7. Drira, H., Ben Amor, B., Srivastava, A., Daoudi, M., Slama, R.: 3D face recognition under expressions, occlusions, and pose variations. IEEE Trans. Pattern Anal. Mach. Intell. **35**(9), 2270–2283 (2013)

8. Forczmański, P., Łabędź, P.: Improving the recognition of occluded faces by means of two-dimensional orthogonal projection into local subspaces. In: Kamel, M., Campilho, A. (eds.) ICIAR 2015. LNCS, vol. 9164, pp. 229–238. Springer, Cham (2015). doi:10.1007/978-3-319-20801-5_25
9. Ganguly, S., Bhattacharjee, D., Nasipuri, M.: Depth based occlusion detection and localization from 3D face image. Int. J. Image Graph. Sig. Process. **7**(5), 20–31 (2015)
10. Gul, S., Farooq, H.: A machine learning approach to detect occluded faces in unconstrained crowd scene. In: 14th International Conference on Cognitive Informatics & Cognitive Computing (ICCI* CC), pp. 149–155. IEEE, July 2015
11. Hassaballah, M., Aly, S.: Face recognition: challenges, achievements and future directions. IET Comput. Vis. **9**(4), 614–626 (2015)
12. Henry, P., Krainin, M., Herbst, E., Ren, X., Fox, D.: RGB-D mapping: using Kinect-style depth cameras for dense 3D modeling of indoor environments. Int. J. Robot. Res. **31**(5), 647–663 (2012)
13. Hsu, G.S.J., Liu, Y.L., Peng, H.C., Wu, P.X.: RGB-D-based face reconstruction and recognition. IEEE Trans. Inf. Forensics Secur. **9**(12), 2110–2118 (2014)
14. Izadi, S., Kim, D., Hilliges, O., Molyneaux, D., Newcombe, R., Kohli, P., Shotton, J., Hodges, S., Freeman, D., Davison, A., Fitzgibbon, A.: KinectFusion: real-time 3D reconstruction and interaction using a moving depth camera. In: Proceedings of the 24th Annual ACM Symposium on User Interface Software and Technology, pp. 559–568. ACM, October 2011
15. Jain, A.K., Nandakumar, K., Ross, A.: 50 years of biometric research: accomplishments, challenges, and opportunities. Pattern Recogn. Lett. **79**, 80–105 (2016)
16. Johnson, R., O'Hara, K., Sellen, A., Cousins, C., Criminisi, A.: Exploring the potential for touchless interaction in image-guided interventional radiology. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pp. 3323–3332. ACM, May 2011
17. KONICA Minolta. http://www.konicaminolta.com/. Accessed 03 Feb 2016
18. Li, X.X., Dai, D.Q., Zhang, X.F., Ren, C.X.: Structured sparse error coding for face recognition with occlusion. IEEE Trans. Image Process. **22**(5), 1889–1900 (2013)
19. Liao, S., Jain, A., Li, S.: A fast and accurate unconstrained face detector. IEEE Trans. Pattern Anal. Mach. Intell. **38**(2), 211–223 (2016)
20. Min, R., Kose, N., Dugelay, J.L.: KinectFaceDB: A Kinect database for face recognition. IEEE Trans. Syst. Man Cybern. Syst. **44**(11), 1534–1548 (2014)
21. Ojala, T., Pietikäinen, M., Mäenpää, T.: Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. IEEE Trans. Pattern Anal. Mach. Intell. **24**(7), 971–87 (2002)
22. Ojala, T., Pietikäinen, M., Mäenpää, T.: A generalized Local Binary Pattern operator for multiresolution gray scale and rotation invariant texture classification. In: ICAPR, vol. 1, pp. 397–406, March 2001
23. Ouloul, M.I., Moutakki, Z., Afdel, K., Amghar, A.: An efficient face recognition using SIFT descriptor in RGB-D images. Int. J. Electr. Comput. Eng. **5**(6), 1227–1233 (2015)
24. Phillips, P.J., Flynn, P.J., Scruggs, T., Bowyer, K.W., Chang, J., Hoffman, K., Marques, J., Min, J., Worek, W.: Overview of the face recognition grand challenge. In: IEEE Computer Society Conference on Computer Vision and Pattern Recognition, CVPR 2005, vol. 1, pp. 947–954, June 2005
25. Saxena, A., Chung, S.H., Ng, A.Y.: 3-d depth reconstruction from a single still image. Int. J. Comput. Vis. **76**(1), 53–69 (2008)

26. Shao, L., Han, J., Xu, D., Shotton, J.: Computer vision for RGB-D sensors: Kinect and its applications. IEEE Trans. Cybern. **43**(5), 1314–1317 (2013)
27. Shotton, J., Sharp, T., Kipman, A., Fitzgibbon, A., Finocchio, M., Blake, A., Cook, M., Moore, R.: Real-time human pose recognition in parts from single depth images. Commun. ACM **56**(1), 116–124 (2013)
28. Srinivasan, A., Balamurugan, V.: Occlusion detection and image restoration in 3D face image. In: TENCON IEEE Region 10 Conference, pp. 1–6, October 2014
29. Sultana, M., Gavrilova, M., Alhajj, R., Yanushkevich, S.: Adaptive multi-stream score fusion for illumination invariant face recognition. In: IEEE Symposium on Computational Intelligence in Biometrics and Identity Management (CIBIM), pp. 94–101, December 2014
30. Sultana, M., Gavrilova, M., Yanushkevich, S.: Fuzzy rule based quality measures for adaptive multimodal biometric fusion at operation time. In: IJCCI, 6th International Conference on Fuzzy Computation Theory and Applications (FCTA), pp. 146–152 (2014)
31. Vapnik, V.N.: An overview of statistical learning theory. IEEE Trans. Neural Networks **10**(5), 988–999 (1999)
32. Viola, P., Jones, M.: Rapid object detection using a boosted cascade of simple features. In: Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, CVPR 2001, vol. 1, p. I-511 (2001)
33. Xia, Y., Zhang, B., Coenen, F.: Face occlusion detection based on multi-task convolution neural network. In: 12th International Conference on Fuzzy Systems and Knowledge Discovery (FSKD), pp. 375–379. IEEE, August 2015
34. Zohra, F.T., Rahman, M.W., Gavrilova, M.: Occlusion detection and localization from kinect depth images. In: International Conference on Cyberworlds (CW), pp. 189–196. IEEE (2016)

Springer

http://www.springer.com/978-3-662-56005-1

Transactions on Computational Science XXX
Special Issue on Cyberworlds and Cybersecurity
Gavrilova, M.; Tan, C.J.K.; Sourin, A. (Eds.)
2017, XIII, 111 p. 59 illus., Softcover
ISBN: 978-3-662-56005-1